Департамент образования и науки города Москвы

Государственное автономное образовательное учреждение

высшего образования города Москвы

«Московский городской педагогический университет»

Институт цифрового образования

Департамент информатики, управления и технологий

Мокрецова Виктория Анатольевна

Тема выпускной квалификационной работы

«Разработка интеллектуальной системы сегментации клиентов и оптимизации программы лояльности на основе многофакторного анализа данных»

**Лабораторная работа 1.1 Установка и настройка ETL-инструмента. Создание конвейеров данных**

Направление подготовки – 38.03.05 «Бизнес-информатика»,

Профиль подготовки – «Аналитика данных и эффективное управление»

(очная форма обучения)

Руководитель:

кандидат технических наук, доцент

Босенко Тимур Муртазович

\_\_\_\_\_

Москва

2025

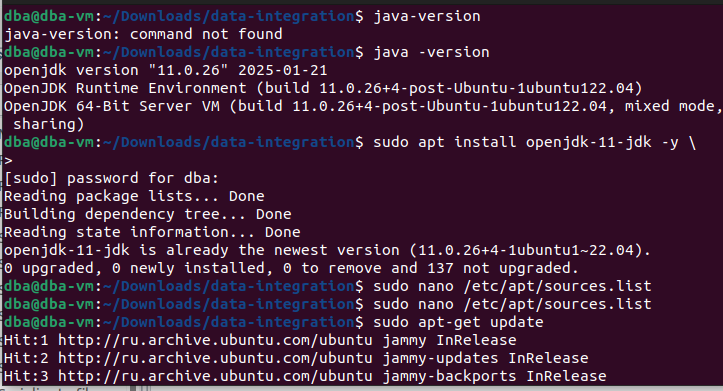
# Цели

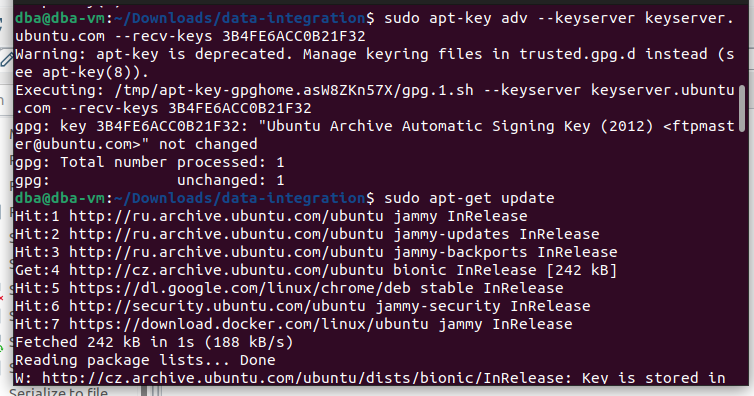
# Научиться открывать пентаго, загружать в него данные, выгружать из него данные в БД.

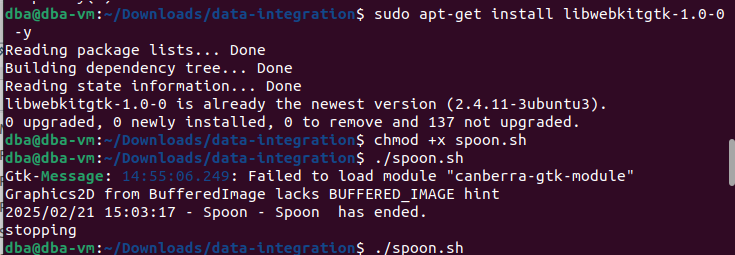
# Задачи:

* Скачать и установить необходимые компоненты в виртуальную среду Virtual Box Ubuntu Linux.
* Загрузить данные из CSV в пентаго,
* Загрузить данные в MSSQL.
* Отобразить результаты.

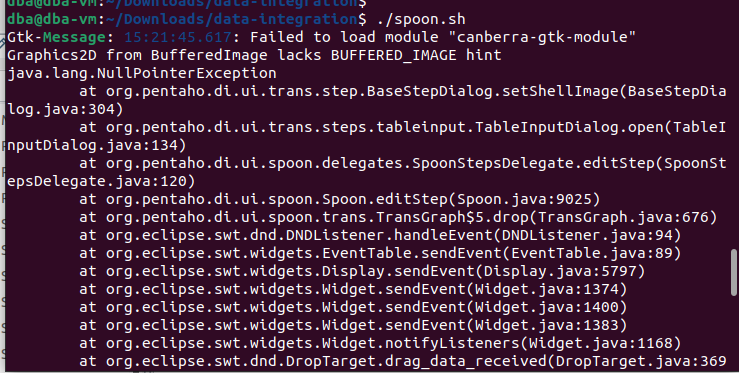
# Установка библиотек.



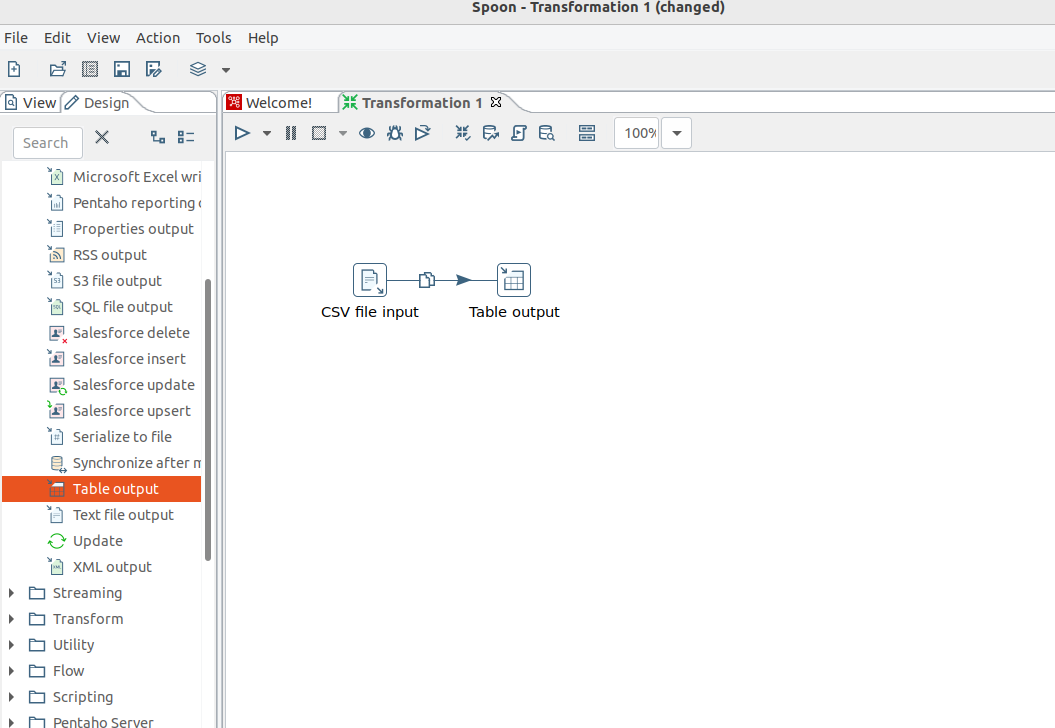




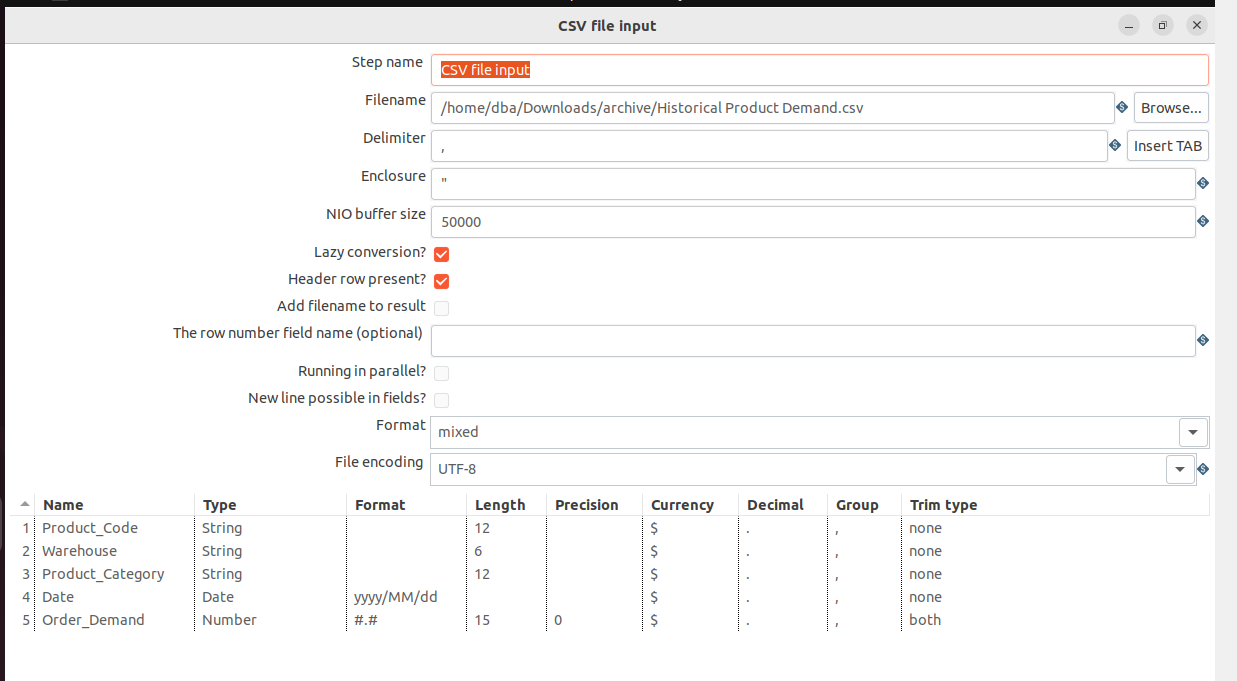
# Запуск пентаго



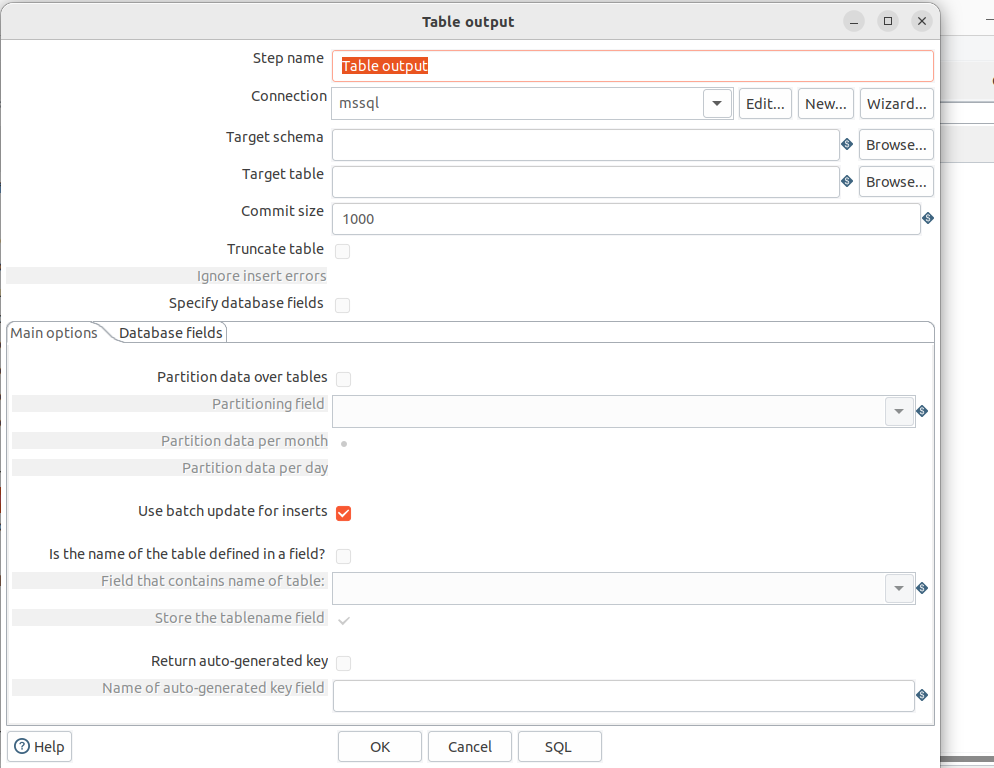
# Добавление входных данных и запись в БД



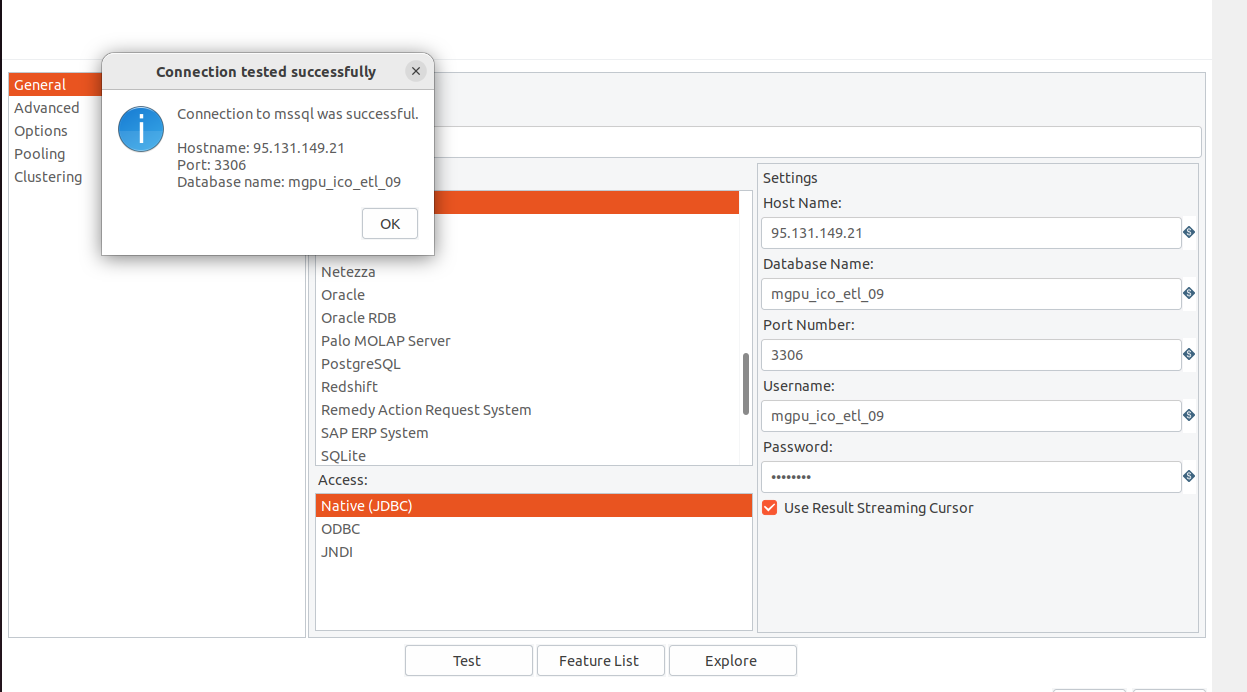
Подключение csv файла, проверка на типы данных



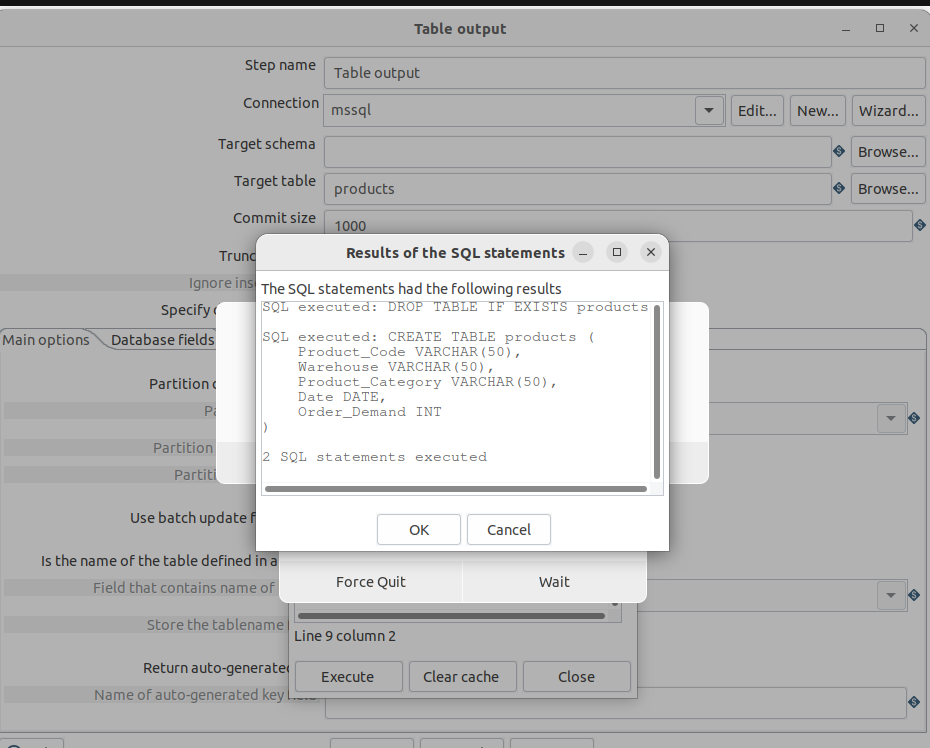
Подключение MSSQl



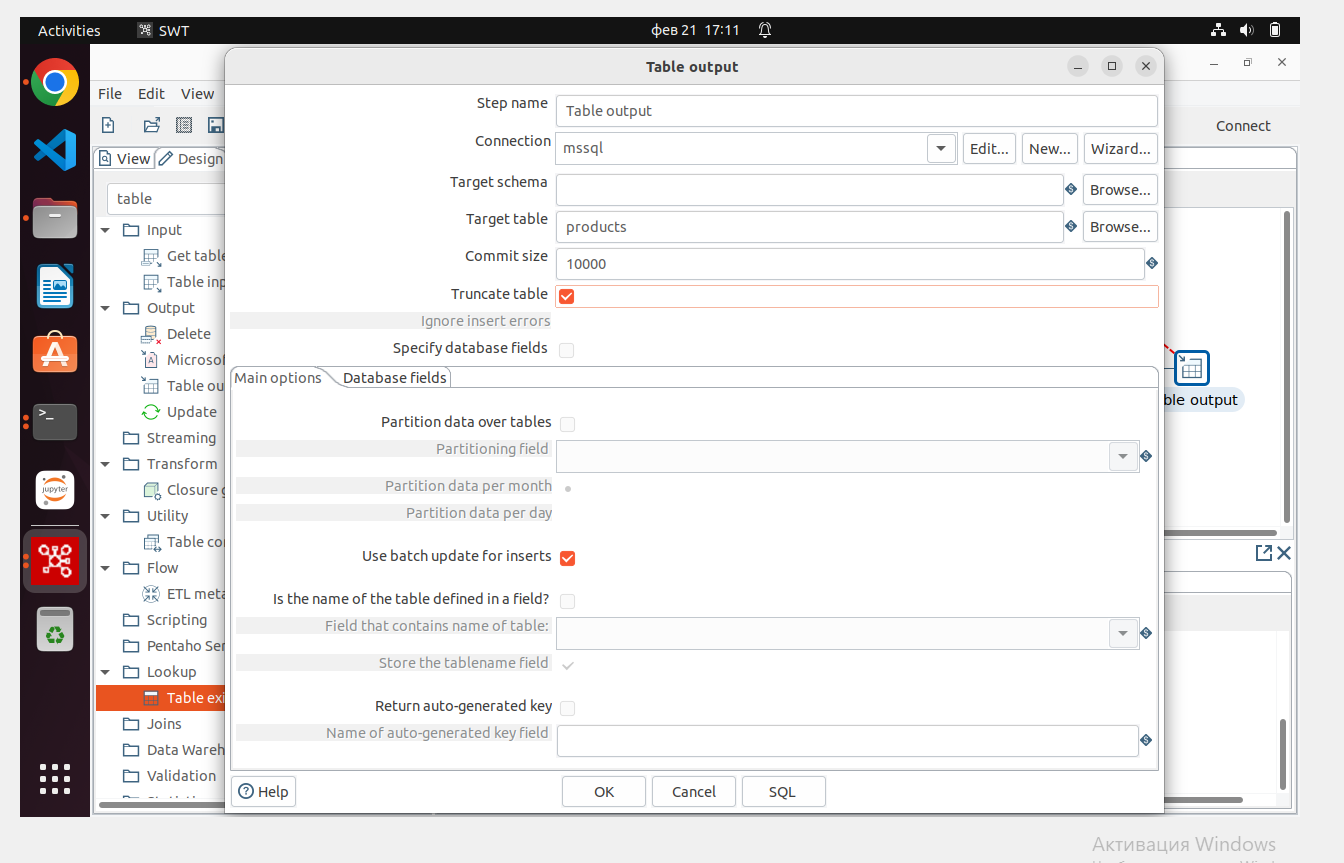
Тестирование подключения к серверу



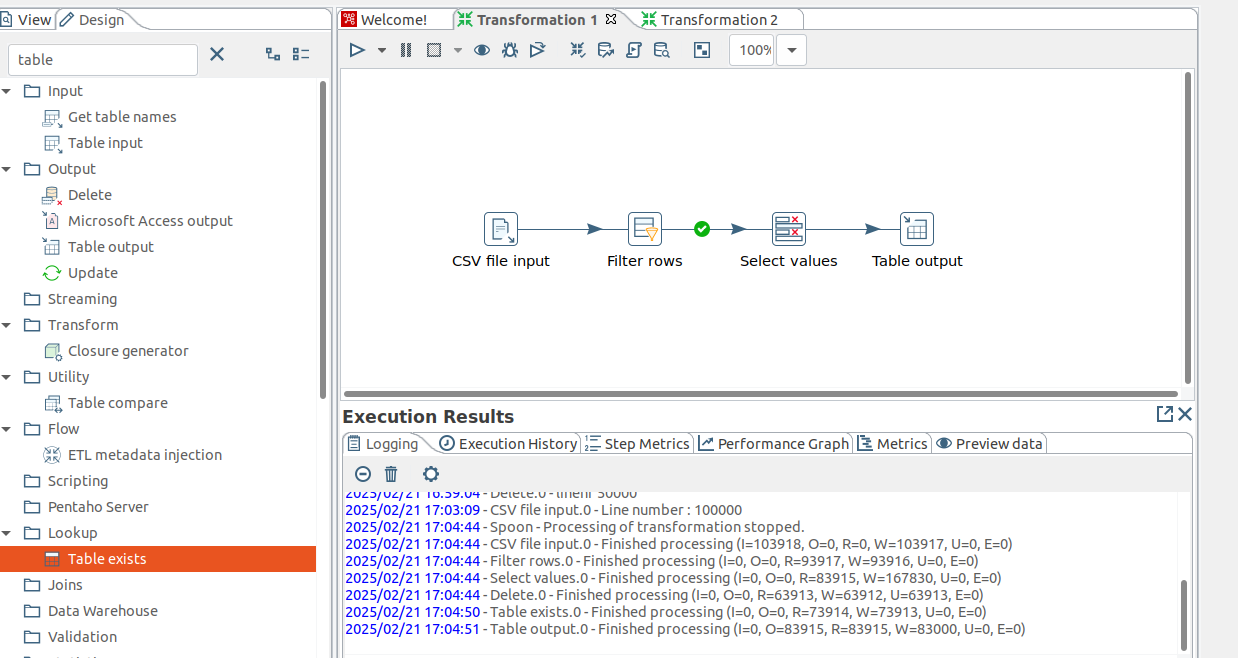
На всякий случай запустила скрипт создания таблицы products.



Добавила две настройки в запись в БД : увеличила размер комита для ускорения процесса записи и добавила удаление данных перед записью из таблицы, так как мы записываем один и тот же файл каждый раз и новых данных не поступает.



Добавила удаление значений NA в поле дата, чтобы избежать ошибки. Этапы: заменила тип данных на входе с «date» на «String», отфильтровала «NA», заменила тип данных обратно на дату.



# Результаты:

